Single Factor ANOVA
(A Full Example)

Lecture Outline

• ANOVA -- what it tells you
• Examine how we might follow up a significant result
• A range of multiple comparison methods
• The David Tiller example
• Reading

Objectives of ANOVA

• To compare two or more means
• To attribute a proportion of total variation to each level of sampling
• To consider simultaneously the effects of two or more variables, and to test for the interactions among them

Dave Tiller’s Study

Tiller’s study aimed to investigate the impact if any of the Thredbo Village on the macro-invertebrate fauna of the adjacent Crackenback River

Note: This is an aim not an objective

Lots of Experimental Design issues here?

Suggestions?
Dave Tiller’s Study

Are there significant differences among the sites in the Crackenback River?

If so, how do the differences relate to the distribution of potential impacts down the river?

If there is an impact of the Village, how far down stream is it manifested?

Note: These are objectives

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>286</td>
<td>325</td>
<td>496</td>
<td>1065</td>
<td>411</td>
<td>207</td>
<td>121</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>669</td>
<td>703</td>
<td>798</td>
<td>1539</td>
<td>1116</td>
<td>104</td>
<td>197</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>142</td>
<td>332</td>
<td>898</td>
<td>1174</td>
<td>681</td>
<td>153</td>
<td>292</td>
<td>468</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>265</td>
<td>640</td>
<td>880</td>
<td>1281</td>
<td>283</td>
<td>208</td>
<td>435</td>
</tr>
</tbody>
</table>

Statistical Outcomes

The differences among the eight sites in abundance of benthic invertebrates (Figure 1) was significant, as demonstrated by ANOVA applied to square root counts ($F = 31.07; df = 7,72; p < 0.0005$)

What Next?

If there is a difference among the means we might well ask:

Where among the means do the differences lie?

The Analysis

(Refer Worked Example)

Multiple comparisons based on the Tukey-Kramer Procedure ($\alpha < 0.05$) revealed that macro-invertebrate abundance was greatest at Site D, immediately downstream of the sewage outflow.
What Next?

We need a procedure that
- avoids compounding errors
- uses all the data in support of comparisons

Multiple Comparison Tests

Tukey-Kramer Procedure

A test based on the T-Test

$$T = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\frac{1}{n} (S_1^2 + S_2^2)}}$$

$$df = 2(n-1)$$

What is our best estimate of the common population variance?

$$MS_{within}$$

The result will be significant if

$$T = \frac{|\bar{Y}_1 - \bar{Y}_2|}{\sqrt{\frac{2}{n} (MS_{within})}} \geq T_{0.05, [2], a(n-1)}$$

$$df = a(n-1)$$
The result will be significant if

\[ |\bar{Y}_1 - \bar{Y}_2| \geq \frac{Q_{\alpha,a,a(n-1)}}{\sqrt{\frac{1}{n} (MS_{\text{within}})}} \]

Tukey’s Honestly Significant Difference

For comparing all means against all other means in an exhaustive set of comparisons

- Uses all the data in each comparison
- Adjusts the level of significance for the number of comparisons
### Student-Newman-Kuels Procedure
A staged method for comparing all means against all other means in an exhaustive set of comparisons
- Uses all the data in each comparison
- Adjusts the level of significance for a restricted number of comparisons

### Dunnett’s Test
For comparing all means against a single control mean
- Uses all the data in each comparison
- Adjusts the level of significance for comparisons involving the control mean only

### Bonferroni Tests
For comparing means in a restricted \textit{a-priori} set of comparisons
- Uses all the data in each comparison
- Adjusts the level of significance for comparisons drawn only from the restricted set

### LSD Tests
For an \textit{a-priori} set of unrelated comparisons
- Uses all the data in each comparison
- Makes no adjustment for the number of comparisons

### My Recommendation
- Exhaustive comparisons? 
  - Tukey-Kramer Procedure
  - All means against a single control? 
    - Dunnett’s Test
  - Restricted \textit{a-priori} set of comparisons? 
    - Bonferroni Correction

### Dave Tiller’s Study
The Follow-up Analysis
(Refer Worked Example)
<table>
<thead>
<tr>
<th>Dave Tiller's Study</th>
<th>Dave Tiller's Study</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statistical Outcomes</strong></td>
<td><strong>Statistical Outcomes</strong></td>
</tr>
<tr>
<td>Site C, immediately above the outflow, and Site E, 1 km below the outflow, were similar in terms of invertebrate abundance, second only to Site D.</td>
<td>The upstream Site A and the downstream Sites F, G and H were not significantly different.</td>
</tr>
<tr>
<td><strong>Conclusions</strong></td>
<td><strong>Conclusions</strong></td>
</tr>
<tr>
<td>The sewage outflow, and to a lesser extent, the Village itself, have a demonstrable impact on macro-invertebrate abundance. Abundance increases, a consequence of a shift to Chironomid larvae at the expense of other forms.</td>
<td>As the macro-invertebrate abundance upstream and &gt; 3 km downstream are similar, it seems likely that the impact of the Village and its sewage outflow had dissipated by the time the stream had flowed 3 km.</td>
</tr>
<tr>
<td><strong>Reading</strong></td>
<td></td>
</tr>
<tr>
<td>Are there significant differences among the sites in the Crackenback River? If so, how do the differences relate to the distribution of potential impacts down the river? If there is an impact of the Village, how far down stream is it manifested?</td>
<td>• Lesson 6, Example 4.1 of Module 4 – The Dave Tiller Study  • Lesson 2 of Module 4 – Multiple Comparisons</td>
</tr>
</tbody>
</table>